

5' NCG CCG GTT CGT CTA CTC TTT CCT TCA GCC GCG TCC TTT CAA CCT TGT CAA CCC
 9 18 27 36 45 54
 63 72 81 90 99 108
 GTC GGC GCG GCC TCT GGT GCA GCG GCG GCT CCT GTT CCT GCC GCA GCT CTC
 117 126 135 144 153 162
 TCC CTT TCT TAC CTC CCC ACC AGA TCC CGG AGA TCG CCC GCC ATG GCT RTA CTT
 T A A R L G T K N A S C L V L A
 171 180 189 198 207 216
 ACT GCG GCC CGG CTC TTG GGA ACC AAG AAT GCA TCT TGT CTT GTC CTT GCA
 T A A R L G T K N A S C L V L A
 225 234 243 252 261 270
 GCC CGG CAT GCC AGT GCT TCC TCC ACG AAT TTG AAA GAC ATA TTG GCT GAC CTG
 A R H A S A S T N L K D I L A D L
 279 288 297 306 315 324
 ATA CCT AAG GAG CAG GCC AGA ATT AAG ACT TTC AGG CAG CAA CAT GGC AAG ACG
 I P K E Q A R I K T F R Q H G K T
 333 342 351 360 369 378
 GTG GTG GGC CAA ATC ACT GTG GAC ATG ATG TAT GGT GGC ATG AGA GGC ATG AAG
 V V G Q I T V D M M Y G G M R G M K

FIGURE 1A

Fig 1B: Amino acid sequence of the protein

GGA	TTC	GTC	TAT	GAA	ACA	TCA	GTT	CTT	GAT	CCT	GAG	GGC	ATC	CGT	TTC	CGA	
G	L	V	Y	E	T	S	V	L	D	P	D	E	G	I	R	F	R
441		450					459			468			477			486	
GGC	TTT	AGT	ATC	CCT	GAA	TGC	CAG	AAA	CTG	CTA	CCC	AAG	GCT	AAG	GGG	GAA	
G	F	S	I	P	E	C	Q	K	L	P	K	A	K	G	G	E	
495		504					513			522			531			540	
GAA	CCC	CTG	CCT	GAG	GGC	TTA	TTT	TGG	CTG	CTG	GTA	ACT	GGA	CAT	ATC	CCA	ACA
E	P	L	P	E	G	L	F	W	L	L	V	T	G	H	I	P	T
549		558					567			576			585			594	
GAG	GAA	CAG	GTA	TCT	TGG	CTC	TCA	AAA	GAG	TGG	GCA	AAG	AGG	GCA	GCT	CTG	CCT
E	E	Q	V	S	W	L	S	K	E	W	A	K	R	A	A	L	P
603		612					621			630			639			648	
TCC	CAT	GTG	GTC	ACC	ATG	CTG	GAC	AAC	TTT	CCC	ACC	AAT	CTA	CAC	CCC	ATG	TCT
S	H	V	V	T	M	L	D	N	F	P	T	N	L	H	P	M	S
657		666					675			684			693			702	
CAG	CTC	AGT	GCA	GCT	GT	ACA	GCC	CTC	AAC	AGT	GAA	AGT	AAC	TTT	GCC	CGA	GCA
Q	L	S	A	A	V	T	A	L	N	S	E	S	N	F	A	R	A
711		720					729			738			747			756	
TAT	GCA	CAG	GGT	ATC	AGC	CGA	ACC	AAG	TAC	TGG	GAG	TTG	ATT	TAT	GAA	GAC	TCT
Y	A	Q	G	I	S	R	T	K	Y	W	E	L	I	Y	E	D	S

FIGURE 1B

ATG	GAT	CTA	ATC	GCA	AAG	CTA	CCT	TGT	GTT	GCA	GCA	AAG	ATC	TAC	CGA	AAT	CTC
M	D	L	I	A	K	L	P	C	V	A	A	K	I	Y	R	N	L
TAC	AGA	GAA	GGC	AGC	GGT	ATT	GGG	GCC	ATT	GAC	TCT	AAC	CTG	GAC	TGG	TCT	CAC
Y	R	E	G	S	G	I	G	A	I	D	S	N	L	D	W	S	H
AAT	TTC	ACC	AAC	ATG	TTA	GGC	TAT	ACT	GAT	CAT	CAG	TTC	ACT	GAG	CTC	ACG	CGC
N	F	T	N	M	L	G	Y	T	D	H	Q	F	T	E	L	T	R
CTG	TAC	CTC	ACC	ATC	CAC	AGT	GAC	CAT	GAG	GGT	GGC	AAT	GTA	AGT	GCC	CAT	ACC
L	Y	L	T	I	H	S	D	H	E	G	G	N	V	S	A	H	T
AGC	CAT	TTG	GTG	GGC	AGT	GCC	CTT	TCC	GAC	CCT	TAC	CTG	TCC	TTT	GCA	GCA	GCC
S	H	L	V	G	S	A	L	S	D	P	Y	L	S	F	A	A	A
ATG	AAC	GGG	CTG	GCA	GGG	CCT	CTC	CAT	GGA	CTG	GCA	AAT	CAG	GAA	GTG	CTT	GTC
M	N	G	L	A	G	P	L	H	G	L	A	N	Q	E	V	L	V
TGG	CTA	ACA	CAG	CTG	CAG	AAG	GAA	GTT	GGC	AAA	GAT	GTG	TCA	GAT	GAG	AAG	TTA
W	L	T	Q	L	Q	K	E	V	G	K	D	V	S	D	E	K	L

FIGURE 1C

CGA	GAC	TAC	ATC	TGG	AAC	ACA	CTC	AAC	TCA	GGA	CGG	GTT	GTC	CCA	GGC	TAT	GGC	1188
R	D	Y	I	W	N	T	L	N	S	G	R	V	V	P	G	Y	G	
1143		1152			1161				1170				1179					1188
CAT	GCA	GTA	CTA	AGG	AAG	ACT	GAT	CCG	CGA	TAT	ACC	TGT	CAG	CGA	GAG	TTT	GCT	
H	A	V	L	R	K	T	D	P	R	Y	T	C	Q	R	E	F	A	1242
1197		1206			1215			1224				1233						1242
CTG	AAA	CAC	CTG	CCT	AAT	GAC	CCC	ATG	TTT	AAG	TTG	GCT	CAG	CTG	TAC	AAG		
L	K	H	L	P	N	D	P	M	F	K	L	V	A	Q	L	Y	K	1296
1251		1260			1269			1278				1287						1296
ATT	GTG	CCC	AAT	GTC	CTC	TTA	GAG	CAG	GGT	AAA	GCC	AAG	AAT	CCT	TGG	CCC	AAT	
I	V	P	N	V	L	L	E	Q	G	K	A	K	N	P	W	P	N	1350
1305		1314			1323			1332				1341						1350
GTA	GAT	GCT	CAC	AGT	GGG	GTG	CTG	CTC	CAG	TAT	TAT	GGC	ATG	ACG	GAG	ATG	AAT	
V	D	A	H	S	G	V	L	Q	Y	Y	G	M	T	E	M	N		
1359		1368			1377			1386				1395						1404
TAC	TAC	ACG	GTC	CTG	TTT	GGG	GTG	TCA	CGA	GCA	TTG	GGT	GTA	CTG	GCA	CAG	CTC	
Y	Y	T	V	L	F	G	V	S	R	A	L	G	V	L	A	Q	L	
1413		1422			1431			1440				1449						1458
ATC	TGG	AGC	CGA	GCC	TTA	GGC	TTC	CCT	CTA	GAA	AGG	CCC	AAG	TCC	ATG	AGC	ACA	
I	W	S	R	A	L	G	F	P	L	E	R	P	K	S	M	S	T	1512
1467		1476			1485			1494				1503						1512

FIGURE 1D

FIGURE 1E

GAG	GGT	CTG	ATG	AAG	TTT	GTG	GAC	TCT	AAG	TCA	GGG	TAA	AAC	TGG	AGA	CTG	GGT	1566
E	G	L	M	K	F	V	D	S	K	S	G	*						
1521	1530	1539	1548	1549	1557	1557												
GAA	AGT	GAC	TAC	CAG	AAA	GTG	AGG	AAG	CCT	AAA	TAA	AAA	GTA	TAC	TTT	TGT	TTC	1620
1575	1584	1593	1602	1602	1611	1611												
AGG	GGG	CCT	TTA	AAG	ACT	TAA	GAT	TAA	ATT	ATA	TCT	GAG	GCA	CTG	ATA	ATA	TGT	1674
1629	1638	1647	1656	1656	1665	1665												
TTG	AGG	TTA	AAA	TAT	AAA	TTA	AGA	CTT	TAA	AAG	ATG	AAA	AAT	GGT	CCC	TTC	TTC	1728
1683	1692	1701	1710	1710	1719	1719												
CCT	AAT	CAG	CTC	CCT	TCC	CCT	GCC	TGG	TAT	GAG	TG	CCC	ATC	ATA	CGC	ATG	GTC	1737
1737	1746	1755	1764	1764	1773	1773												
CTG	GAG	GAT	GAC	CAG	GAC	TAA	TGC	ATG	TGG	TAT	GAG	TAG	GTT	TGG	CCC	CCT	CAC	1791
1845	1854	1863	1872	1872	1881	1881												
TAT	CTC	TAG	AGT	GAG	AAT	CTG	GCT	CCT	GTT	TCC	ATG	GGT	CAA	AGC	CGG	TTG	CAG	1890

FIGURE 1F

1899	1908	1917	1926	1935	1944
AGA	ATC	TGT	AGT	CAC	TGT
	GCA	TTC	TGA	GCT	TTA
	CCC	TTC	TTC	TCT	TCT
	TTC	TGT	CAT	AGG	AAT
			CAT	GTT	GGA
				TAG	TCA
1953	1962	1971	1980	1989	1998
AGC	AAA	CCA	GGA	CTC	TGC
	TTC	CCC	TCC	TTC	TTC
			CAT	CAT	CAT
				AGG	TAA
				CCT	GCC
2007	2016	2025	2034	2043	2052
GCT	GTA	CCA	AGC	CCC	TTC
				GCC	CTC
				CAC	CAC
				AAA	CAC
				CTC	CTC
				CTA	GCA
				AGA	AGA
2061	2070	2079	2088	2097	2106
CCT	GTT	GGT	TAG	CTG	GAC
			ATG	CTT	TGG
				CAA	TTT
				TTT	TAT
				TAT	ACT
				ACC	AAG
				TGA	TGA
				CCA	CCA
2115	2124				
TAT	TGG	CAT	GGC	ATT	TTT
				TGG	TGA
				TG	3'

1	M A L L T A A A R L L G T K N A S C L V	2171653
1	M A L L T A A A R L F G A K N A S C L V	GI 164419
21	L A A R H A S A S S T N L K D I L A D L	2171653
21	L A A R H A S A S S T N L K D I L A D L	GI 164419
41	I P K E Q A R I K T F R Q Q H G K T V V	2171653
41	I P K E Q A R I K T F R Q Q H G N T V V	GI 164419
61	G Q I T V D M M Y G G M R G M K G L V Y	2171653
61	G Q I T V D M M Y G G M R G M K G L V Y	GI 164419
81	E T S V L D P D E G I R F R G F S I P E	2171653
81	E T S V L D P D E G I R F R G Y S I P E	GI 164419
101	C Q K L L P K A K G G E E P L P E G L F	2171653
101	C Q K M L P K A K G G E E P L P E G L F	GI 164419
121	W L L V T G H I P T E E Q V S W L S K E	2171653
121	W L L V T G Q I P T E E Q V S W L S K E	GI 164419
141	W A K R A A L P S H V V T M L D N F P T	2171653
141	W A K R A A L P S H V V T M L D N F P T	GI 164419
161	N L H P M S Q L S A A V T A L N S E S N	2171653
161	N L H P M S Q L S A A I T A L N S E S N	GI 164419
181	F A R A Y A Q G I S R T K Y W E L I Y E	2171653
181	F A R A Y A E G I H R T K Y W E L I Y E	GI 164419
201	D S M D L I A K L P C V A A K I Y R N L	2171653
201	D C M D L I A K L P C V A A K I Y R N L	GI 164419
221	Y R E G S G I G A I D S N L D W S H N F	2171653
221	Y R E G S S I G A I D S K L D W S H N F	GI 164419

FIGURE 2A

241	T N M L G Y T D H Q F T E L T R L Y L T	2171653
241	T N M L G Y T D A Q F T E L M R L Y L T	GI 164419
261	I H S D H E G G N V S A H T S H L V G S	2171653
261	I H S D H E G G N V S A H T S H L V G S	GI 164419
281	A L S D P Y L S F A A A M N G L A G P L	2171653
281	A L S D P Y L S F A A A M N G L A G P L	GI 164419
301	H G L A N Q E V L V W L T Q L Q K E V G	2171653
301	H G L A N Q E V L V W L T Q L Q K E V G	GI 164419
321	K D V S D E K L R D Y I W N T L N S G R	2171653
321	K D V S D E K L R D Y I W N T L N S G R	GI 164419
341	V V P G Y G H A V L R K T D P R Y T C Q	2171653
341	V V P G Y G H A V L R K T D P R Y T C Q	GI 164419
361	R E F A L K H L P N D P M F K L V A Q L	2171653
361	R E F A L K H L P H D P M F K L V A Q L	GI 164419
381	Y K I V P N V L L E Q G K A K N P W P N	2171653
381	Y K I V P N V L L E Q G K A K N P W P N	GI 164419
401	V D A H S G V L L Q Y Y G M T E M N Y Y	2171653
401	V D A H S G V L L Q Y Y G M T E M N Y Y	GI 164419
421	T V L F G V S R A L G V L A Q L I W S R	2171653
421	T V L F G V S R A L G V L A Q L I W S R	GI 164419
441	A L G F P L E R P K S M S T E G L M K F	2171653
441	A L G F P L E R P K S M S T D G L I K L	GI 164419
461	V D S K S G	2171653
461	V D S K	GI 164419

FIGURE 2B